

С. Д. Штовба, д. т. н., доц.; О. В. Штовба, к. е. н.

SH-ІНДЕКС – НОВА ДРОБОВА МОДИФІКАЦІЯ ІНДЕКСА ХІРША

Запропоновано новий наукометричний показник – Sh-індекс. В Sh-індексі ціла частина є звичайним індексом Хірша, а дробова показує, наскільки науковець наблизився до наступного значення індекса Хірша. Дробову частину розраховують як частку вже опублікованих статей, які потрібні для отримання наступного значення індекса Хірша.

Ключові слова: наукометрія, індекс цитування, індекс Хірша, дробова модифікація, Sh-індекс.

Вступ

Сьогодні для оцінювання продуктивності діяльності вчених спільно з експертними висновками все частіше використовують і наукометричні показники. Інтерес до наукометричних показників каталізували роботи радянського вченого В. В. Налімова, найвідомішою серед яких є монографія [1]. Спочатку основним наукометричним показником була кількість друкованих праць ученого – загальна або за окремими типами: монографії, статті, тези, публікації в спеціальних виданнях, що увійшли до списку ВАК України, внесені в електронні бази Web of Science або Scopus чи проіндексовані Google Scholar тощо. Для штучного покращення цього показника використовують такі типові прийоми, як: дроблення результатів для друку в різних виданнях або навіть дублювання опису одних і тих самих матеріалів у статтях з різними назвами. Таким чином, гонитва за кількістю публікацій часто знижує якість наукових робіт.

Другим популярним наукометричним показником став індекс цитування – кількість публікацій з посиланням на автора, роботи якого аналізуються. Індекс цитування характеризує реакцію наукової спільноти на публікації автора, тому його доцільно використовувати для оцінювання результатів фундаментальних досліджень. Як правило, слабкі роботи не цитують, за винятком особливих стосунків між авторами. Різні модифікації індекса цитування очищують його від «шумів», наприклад, виключають самоцитування чи цитування співавторами або вимагають наявності в одній базі даних і статті, і робіт з посиланням на неї. Щоб потрапити до цієї бази даних статті мають відповідати певним вимогам. Але і цей показник не захищає від розпорошення результатів дослідження по дрібних статтях. Можливо, великий індекс цитування забезпечує лише одна стаття, яку науковець написав багато років тому і після того нічого суттєвого не публікував. Тому за одним лише індексом цитування важко достовірно оцінити рівень продуктивності науковця, який важливий для прийняття на роботу, фінансування досліджень тощо.

Щоб виявити дослідників, які пишуть багато і якісно, у 2005 р. фізик Х. Хірш запропонував новий показник – індекс Хірша [2]. Індекс Хірша або h-індекс – це найбільше ціле число h , яке вказує, що автор опублікував щонайменше h статей, кожна з яких процитована хоча б h разів. Ці h статей отримали назву ядро Хірша або h-ядро. Щоб потрапити до ядра Хірша, статтю мають процитувати принаймні h разів, тому не слід дробити результати по кільком публікаціям. Простота розрахунків та нечутливість до типових прийомів штучного покращення вищезгаданих показників стрімко зробили індекс Хірша популярним наукометричним індикатором.

Розглянемо двох авторів, які за мінімальних вимог досягли сусідніх значень індекса Хірша h та $h+1$. Кількість цитувань цих авторів відрізняється суттєво, а саме – на величину $(h+1)^2 - h^2 = 2h+1$. Тому за індексом Хірша можна просто кластеризувати науковців,

наприклад, на групи з однаковими його значеннями. Але цілочисельність цього показника не дозволяє ранжувати авторів у межах групи з однаковими значеннями індекса Хірша. **Метою статті** є розробка нового наукометричного показника, що успадкує всі переваги індекса Хірша та додатково дозволить здійснити таке ранжування. Новий показник є дробовою модифікацією індекса Хірша. У ній ціла частина є звичайним індексом Хірша, а дробова – показує, наскільки наблизився науковець до наступного значення індекса Хірша. Відповідно до наукометричних традицій новий показник назвемо Sh-індексом.

Формалізована постановка завдання

Позначимо кількість публікацій деякого автора через N . Кількість цитувань його публікацій упорядкуємо за спаданням і представимо таким вектором:

$$C = (c_1, c_2, \dots, c_N), \quad (1)$$

де c_i – кількість цитувань i -ої публікації, $c_1 \geq c_2 \geq \dots \geq c_N$, $i = \overline{1, N}$.

Завдання дослідження полягає в знаходженні функціонального відображення

$$C \rightarrow Sh_index \in [h(C), h(C) + 1], \quad (2)$$

де Sh_index – дробова модифікація індекса Хірша (Sh-індекс);

$$h(C) = \max_{\substack{i=1, N \\ i \geq c_i}}(i) \text{ – індекс Хірша.}$$

На відображення (2) накладемо такі обмеження:

(а) якщо індекс Хірша $h(C)$ набрано за мінімально можливою кількістю цитувань, тобто коли $(c_1 = h, c_2 = h, \dots, c_h = h)$, тоді $Sh_index(C) = h(C)$;

(б) чим більша дробова частина Sh_index , тим ближче автор до збільшення свого індекса Хірша.

Новий Sh-індекс

Новий наукометричний Sh-індекс для вектора цитувань (1) визначимо так:

$$Sh_index(C) = h(C) + \Delta, \quad (3)$$

де $\Delta = \frac{1}{h+1} \cdot \max_{\substack{j=1, h \\ c_j > h}}(j)$ – дробова частина Sh-індекса, яка показує частку наповненості

поточного h -ядра статтями, які за числом цитувань входять і до наступного ядра.

Для зручності розрахунків дробовий складник (3) перепишемо в такому вигляді:

$$\Delta = \frac{\sum_{j=1, h} r_j}{h+1}, \quad (4)$$

де $r_j = \begin{cases} 1, & \text{якщо } c_j > h \\ 0, & \text{інакше} \end{cases}$ – індикатор належності статті до наступного ядра Хірша.

Приклад. Вектори цитувань трьох авторів наведено в табл. 1. У ній усі автори мають однаковий індекс Хірша, який дорівнює $h = 4$. Щоб його підвищити на 1, потрібно 5 статей, кожна з яких має бути процитована щонайменше 5 разів. Автор А має 4 таких статті, автор Б – 3 таких статті, а автор С – 0. Відповідно оцінки авторів за запропонованим Sh-індексом дорівнюють:

$$sh_index(\text{Автор А}) = 4 + \frac{4}{5} = 4.8;$$

$$sh_index(\text{Автор B}) = 4 + \frac{3}{5} = 4.6;$$

$$sh_index(\text{Автор B}) = 4 + \frac{0}{5} = 4.$$

Таблиця 1

Вектори цитувань

| Автори | Номер статті | | | | | | | |
|---------|--------------|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Автор А | 20 | 9 | 7 | 5 | 4 | 2 | 1 | 0 |
| Автор Б | 6 | 5 | 5 | 4 | 1 | 0 | 0 | 0 |
| Автор В | 4 | 4 | 4 | 4 | 0 | 0 | 0 | 0 |

Порівняння з конкурентними підходами

Найчастіше для ранжування науковців з однаковими індексами Хірша застосовують додаткові показники, найпопулярніші з яких зведено в табл. 2. Перевага запропонованого показника полягає в тому, що він узагальнює індекс Хірша, тоді як показники з табл. 2 використовують спільно з індексом Хірша. Переваги та недоліки цих додаткових наукометричних показників проаналізовано в багатьох статтях, наприклад в [9, 10], тому ми на них не зупинятимемося.

Таблиця 2

Додаткові наукометричні показники

| Показник | Опис | Відношення | Джерело |
|-----------|---|------------------------|---------|
| g-індекс | Максимальне число g найпопулярніших статей, які спільно отримали не менше g^2 цитувань. Враховує перевищення сумарного цитувань ядра Хірша над мінімально встановленими вимогами. | $h \leq g$ | [3] |
| hg-індекс | Середнє геометричне h-індекса та g-індекса: $hg = \sqrt{h \cdot g}$. | $h \leq hg \leq g$ | [4] |
| e-індекс | Квадратний корінь надлишкового цитування ядра Хірша h^2 : $e = \sqrt{\sum_{j=1, h} c_j - h^2}$. | – | [5] |
| A-індекс | Середнє число цитувань ядра Хірша: $A = \frac{1}{h} \sum_{j=1, h} c_j$. | $R = \sqrt{h \cdot A}$ | [6] |
| R-індекс | Квадратний корінь із сумарного цитування ядра Хірша: $R = \sqrt{\sum_{j=1, h} c_j}$. | $R = \sqrt{h \cdot A}$ | [7] |
| m-індекс | Медіана кількості цитувань ядра Хірша. | $h \leq m$ | [8] |

Найближчим аналогом запропонованого Sh-індекса є h_{rat} -індекс [11]:

$$h_{rat}(C) = h(C) + 1 - \frac{n}{2h(C) + 1}, \quad (5)$$

де $n = (h(C) + 1 - c_{j+1}) + \sum_{j=1, h} (1 - r_j)$ – мінімальна кількість цитувань для отримання наступного значення індекса Хірша;

$2h(C) + 1$ – теоретична мінімальна кількість цитувань для отримання наступного значення індекса Хірша ($h + 1$) за найгірших стартових умов, коли

$(c_1 = h, c_2 = h, \dots, c_h = h, c_{h+1} = 0)$.

Порівнюючи (3) та (5), бачимо, що цілі частини обох показників тотожні й дорівнюють індексу Хірша. Дробові частини відрізняються через різні принципи до інтерполяції відрізка $[h, h+1]$. У (5) цей одиничний відрізок розбивають на $2h(C)+1$ рівні частини відповідно до мінімально потрібної кількості цитувань. Далі до h – кількості публікацій з поточного ядра Хірша – додається частка виконання плану цитувань для утворення наступного ядра Хірша. Відповідно в h_{rat} існує протиріччя між складниками тому, що один доданок відповідає кількості публікацій, а другий – формується за кількістю цитувань. У запропонованому підході цей недолік відсутній, тому що і ціла, і дробова складники Sh-індексу (3) сформовані за кількістю статей. У (3) ціла частина дорівнює кількості статей з поточного ядра Хірша, а дробова – частці заповненого статями наступного ядра. Іншими словами, дробову частину у Sh-індексі можна інтерпретувати як частку виконання плану за статтями для отримання наступного значення індекса Хірша.

Висновки

Запропоновано новий наукометричний показник – Sh-індекс. Цей показник є дробовою модифікацією одного із найпопулярніших наукометричних показників – індекса Хірша. У Sh-індексі ціла частина є звичайним індексом Хірша, а дробова показує, наскільки наблизився науковець до наступного значення індекса Хірша. Дробову частину розраховують як частку вже опублікованих статей, які потрібні для отримання наступного значення індекса Хірша. Успадковуючи всі переваги індекса Хірша, новий показник додатково дозволяє ранжувати авторів з однаковими індексами Хірша. Крім того, за новим показником можна відслідковувати прогрес науковця в покращенні свого індекса Хірша за певний проміжок часу.

СПИСОК ЛІТЕРАТУРИ

1. Налимов В. В. Наукометрия. Изучения развития науки как информационного процесса / В. В. Налимов, З. М. Мульченко. – М.: Наука. – 1969. – 192 с.
2. Hirsch J. E. An index to quantify an individual's scientific research output / J. E Hirsch // Proceedings of the National Academy of Sciences of the USA. – 2005. – Vol. 102, № 46. – P. 16569 – 16572.
3. Egghe L. Theory and practice of the g-index / L. Egghe // Scientometrics. – 2006. – Vol. 69, № 1. – P. 131 – 152.
4. Alonso S. Hg-index: a new index to characterize the scientific output of researchers based on the hand g-indices / S. Alonso, F. Cabrerizo, E. Herrera-Viedma, F. Herrera // Scientometrics. – 2010. – Vol. 82, № 2. – P. 391 – 400.
5. The e-index, complementing the h-index for excess citations [Електронний ресурс] / Zhang C.-T. // PLoS ONE. – Vol. 4, №5. – 2009. Режим доступу: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0005429>.
6. New developments related to the Hirsch index [Електронний ресурс] / Rousseau R. // Science Focus. – Vol. 1, № 4. – 2006. – P. 23 – 25. Режим доступу: <http://eprints.rclis.org/6376/>.
7. Liang B. J. L. The R- and AR-indices: complementing the h-index / B. J. L. Liang, R. Rousseau, L. Egghe // Chinese Science Bulletin. – 2007. – Vol. 52, № 6. – P. 855 – 863.
8. Bornmann L. Are there better indices for evaluation purposes than the h-index? A comparison of nine different variants of the h-index using data from biomedicine / L. Bornmann, R. Mutz, H. Daniel // Journal of the American Society for Information Science and Technology. – 2008. – Vol. 59, № 5. – P. 830 – 837.
9. Alonso S. H-index: a review focused in its variants, computation and standardization for different scientific fields / S. Alonso, F. J. Cabrerizo, E. Herrera-Viedma, F. Herrera // Journal of Informetrics. – 2009. – Vol. 3, № 4. – P. 273 – 289.
10. Egghe L. The Hirsch-index and related impact measures / L. Egghe // Annual Review of Information Science and Technology. – 2010. – Vol. 44. – P. 65 – 114.
11. Guns R. Real and rational variants of the h-index and the g-index / R. Guns, R. Rousseau // Journal of Informetrics. – 2009. – Vol 3, № 1 – P. 64 – 71.

Штовба Сергій Дмитрович – д. т. н., професор кафедри комп'ютерних систем управління.

Штовба Олена Валеріївна – к. н. т., доцент кафедри менеджменту та моделювання в економіці.
Вінницький національний технічний університет.